DOCUMENT RESUME

ED 316 556 TM 014 506

AUTHOR	Sawilowsky, Shlomo S.; Markman, Barry S.
TITLE	Another Look at the Power of Meta-Analysis in the Solomon Four-Group Design.
PUB DATE	88
NOTE	13p.; Paper comments on "Statistical Treatment of the Solomon Four-Group Design: A Meta-Analytic Approach" by M. C. W. Braver and S. L. Braver (1988).
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Equations (Mathematics); Mathematical Models; *Meta Analysis; *Statistical Significance
IDENTIFIERS	*Solomon Four Group Design

ABSTRACT

This paper demonstrates that a meta-analysis technique applied to the Solomon Four-Group Design (SFGD) can fail to find significance even though an earlier "weaker" test may have found significance. The meta-analysis technique was promoted by Braver and Braver as the most powerful single test for analyzing data from an SFGD. They demonstrated that customary analysis on fabricated data sometimes failed to find significance, while meta-analysis found significance. The current paper counters that meta-analysis is a five-step process, and that fabricated data sets can easily be constructed to show that customary tests can find significance, whereas meta-analysis does not. Statements about the efficacy and power of meta-analysis should be made more cautiously. (TJH)

Another Look At The Power Of Meta-Analysis In The Solomon Four-Group Design

SHLOMO S. SAWILOWSKY

Educational Evaluation and Research
College of Education
Wayne State University

BARRY S. MARKMAN

Educational Psychology
College of Education
Wayne State University

Another Look At The Power Of Meta-Analysis In The
Solomon Four-Group Design

Braver and Braver (1988) promoted a meta-analysis technique as the most powerful single test for analyzing data from a Solomon Four-Group Design. Customary analysis on fabricated data failed to find significance, but the meta-analysis found significance. The current paper counters that meta analysis is a five-step process, and further, fabricated data sets can easily be constructed to show that customary tests can find significance, whereas the meta-analysis does not. Statements about the efficacy and power of meta-analysis should be made more cautiously.

Another Look At The Power Of Meta-Analysis In The

Solomon Four-Group Design

Recently, Braver and Braver (1988) made a strong case in *Psychological Bulletin* for reconsidering the under-used Solomon Four-Group Design (Table 1). They pointed out that perhaps its most serious drawback is the inability of the researcher to pinpoint the precise statistical treatment of the design. The data analysis solution proposed by Braver and Braver is a meta-analysis technique by Stouffer (see Braver and Braver for references).

They provided a flowchart to help visualize the five step process that leads to the meta-analysis. In brief, it begins with a 2x2 ANOVA (referred to as Test A) on O_2, O_4, O_5, and O_6, the post test scores for the four groups. Failing to find significance, the flowchart proceeds to a main effects test on the experimental vs. control effect (Test D). If significant results are still not found, the process continues with either an ANCOVA (Test E) on O_2 and O_4 with O_1 and O_3 as covariates, a gain score analysis (Test F), or a repeated measures ANOVA (Test G), as appropriate. If the result at this phase is not significant, the flowchart moves on to a t test on O_5 and O_6 (Test H). Only when a nonsignificant result again appears does the flowchart bring the researcher to the meta-analysis (Test I).

It should be noted that if significance is found at any of the four steps prior to the meta-analysis, Braver and Braver advised, "further testing may cease." Thus, meta-analysis is only performed when the customary statistical treatment fails to indicate a significant treatment effect.

The formula for the meta-analysis is

$$Z_{meta} = \Sigma_i Z_{pi} / \sqrt{k} ,$$

where Z_{pi} is the Z value of the one-tailed p value of the ith statistical test and k=number of tests.

Braver and Braver said, "Test I allows full use of all data and thus is the most powerful single test of the treatment effect available." To demonstrate the power of Test I they constructed a fabricated data set. The series of four customary tests failed to find significance. Then, they performed a meta-analysis on the last two tests (E and H), finding a significant result of p=.040. On this basis they concluded, "These hypothetical data show the superior power of the meta-analytic technique, because none of the customary analyses reached significance but the meta-analysis did."

There are several concerns with Braver and Braver's conclusion. A minor objection the applied researcher may have is with the view of a five step meta-analysis as a single test, especially since it cannot be done at all until at least tests E and H have been performed. Typically in meta-analysis the original studies were previously performed, requiring but a single method applied to those results. Second, meta-analysis techniques are usually employed with apparently diverse or seemingly conflicting results, as opposed to the situation when studies all reached the same statistical conclusion. Although in theory meta-analysis could be used when all of the studies are barely nonsignificant, we have never seen such a situation published. Braver and Braver's prescription for meta-analysis is restricted to the when all previous analyses reached the same (nonsignificant) conclusion.

A more serious concern is promoting the meta-analysis technique as the most powerful test for the Solomon Four-Group Design. Although a most powerful test should find significance (if it exists) when weaker tests cannot, certainly where a weaker test finds significance, so too should a most powerful test.

This paper will demonstrate that the meta-analysis technique can fail to find significance even though an earlier "weaker" test found significance. This is particularly disquieting, since Braver and Braver advised to cease testing at any step where significant results were obtained, thereby never allowing for this possibility to surface.

Table 2 contains a set of fabricated data. Test E (ANCOVA) and H (t) required by the meta-analysis, and Test I (meta-analysis) will be performed on the data. Table 3 reports summary information for the data. Table 4 indicates that Test E (ANCOVA) was not significant, with p=.934. Table 5 shows the result for Test H (t), t = 2.07, with p = .048, a significant result.

[Place Tables 1-5 about here]

The Z_{meta} is

$$Z_{meta} = \frac{1.98+.08}{\sqrt{2}}$$

$$= 1.4567$$

which yields a p of .072. As Braver and Braver said, it "is substantial but not quite significant by conventional standards".

This set of fabricated data is not an exception to the rule. Any number of data sets can be constructed where one of the earlier tests is significant and the meta-analysis is not, by solving for "pb" when

$$\frac{Z_{pa} + Z_{pb}}{\sqrt{2}} < 1.65 \quad ,$$

where Z_{pa} refers to the Z associated with the p value of Test a, Z_{pb} is the Z associated with the p value of Test b, and $pa \leq .05$. So too, any number of data sets can be fabricated to show results like Braver and Braver obtained.

If it is prudent to argue, as we do, that the use of a single fabricated data set is insufficient evidence of meta-analysis being the most powerful technique in the Solomon Four-Group Design, is it equally prudent to say that a single fabricated data set is insufficient in showing the opposite. Braver and Braver's innovative use of meta-analysis in the Solomon Four-Group Design should reawaken the interest of the researcher concerned with external validity issues. However, its efficacy and power in comparison to traditional analyses remains to be seen, and we advise using the technique with caution.

References

Braver, M. C. W., and Braver, S. L. (1988). Statistical treatment of the Solomon four-group design: a meta-analytic approach. *Psychological Bulletin*, **104**, No. 1, 150-154.

Table 1.

Solomon Four-Group Design.

	Group	Pretest	Treatment	Posttest
R	1	O_1	X_1	O_2
R	2	O_3		O_4
R	3		X_2	O_5
R	4			O_6

Note. 0 = outcome measure; X = treatment; R = randomization

Table 2.

Fabricated Data For Solomon Four-Group Design.

Observation					
O_1	O_2	O_3	O_4	O_5	O_6
64	54	63	54	56	50
64	59	60	62	64	50
66	55	59	62	64	51
64	56	58	51	55	51
64	55	59	57	59	51
63	65	65	56	57	61
61	64	62	56	57	60
58	64	58	56	59	61
58	65	60	61	62	62
59	65	58	70	71	60
57	64	66	70	71	60
62	65	63	67	68	62
65	68	65	75	76	70
66	69	58	75	78	69

Table 3.

Summary Statistics for Hypothetical Data.

Group	Pretest *M*	Pretest Variance	Correlation between pre- and posttest	Postest *M*	Postest Variance
1	62.2	9.6		62.0	26.1
			-.33		
2	61.0	8.6		62.3	62.2
			.18		
3				64.1	57.8
4				58.4	46.1

Note. *N* = 14 *per group.*

Table 4.

Test E: ANCOVA on O_2 and O_4 with O_1 and O_3 As Covariates

Source	MS	df	F	p
Treatment vs. Not	.321	1	.007	.934
Error	1200.04	25		

Table 5.

Test H: t on O_5 and O_6.

t	df	p
2.07	26	.048